



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>6</sup> :</b> <b>C12N 15/12, 15/82, C07K 14/81, A01N 63/02, A01H 5/00</b>	<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 95/24479</b> <b>(43) International Publication Date:</b> 14 September 1995 (14.09.95)
<b>(21) International Application Number:</b> PCT/EP95/00881 <b>(22) International Filing Date:</b> 9 March 1995 (09.03.95) <b>(30) Priority Data:</b> 08/208,571                      10 March 1994 (10.03.94)                      US <b>(71) Applicant (for all designated States except AT DE US):</b> SANDOZ LTD. [CH/CH]; Lichtstrasse 35, CH-4002 Basle (CH). <b>(71) Applicant (for DE only):</b> SANDOZ-PATENT-GMBH [DE/DE]; Humboldtstrasse 3, D-79539 Lörrach (DE). <b>(71) Applicant (for AT only):</b> SANDOZ-ERFINDUNGEN VERWALTUNGSGESELLSCHAFT M.B.H. [AT/AT]; Brunner Strasse 59, A-1230 Vienna (AT). <b>(72) Inventors; and</b> <b>(75) Inventors/Applicants (for US only):</b> BRUNKE, Karen, Joan [US/US]; 11 Somerset Court, Belmont, CA 94002 (US). CHAN, Victor, Jack [US/US]; 566 58th Street, Oakland, CA 94609 (US). DELUCA-FLAHERTY, Camille, Renee [US/US]; 4159 Park Boulevard, Palo Alto, CA 94306 (US). SCARAFIA, Liliana, Estela, Cristina [AR/US]; 1031 Varsity Court, Mountain View, CA 94040 (US).		<b>(74) Common Representative:</b> SANDOZ LTD.; Patents & Trade-marks Div., Lichtstrasse 35, CH-4002 Basle (CH). <b>(81) Designated States:</b> AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, ES, FI, GB, GE, HU, JP, KE, KG, KP, KR, KZ, LK, LT, LU, LV, MD, MG, MN, MW, MX, NL, NO, NZ, PL, PT, RO, RU, SD, SE, SI, SK, TJ, TT, UA, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ, UG).  <b>Published</b> <i>With international search report.</i>
<b>(54) Title:</b> NOVEL THIOL PROTEASE INHIBITOR  <b>(57) Abstract</b> <p>A novel inhibitor peptide (designated virgiferin) of thiol proteases is isolated from <i>Diabrotica virgifera</i>. The DNA encoding virgiferin and modified virgiferin peptides are claimed. These sequences may be cloned into vectors and used to transform plants thus conferring reduced susceptibility to damage by plant pests that have thiol proteases as digestive enzymes including insects and nematodes and particularly Coleopteran insects.</p>		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

## NOVEL THIOL PROTEASE INHIBITOR

This invention is concerned with novel inhibitor peptides of proteases, nucleic acid sequences encoding these peptides, incorporation of protease inhibitor genes into the genome of a plant and the expression of the inhibitor gene or genes in plants wherein said plants have reduced susceptibility to damage caused by plant pests including insects and nematodes. The invention is further directed to biotechnological methods for producing said inhibitor peptides, insecticidal compositions containing said inhibitor peptides and the use of said peptides in combatting or reducing insect damage to plants.

Certain plant pests particularly *Coleopteran* insects and those insects of the genus *Diabrotica* cause tremendous damage to crop plants. In the past and currently, chemical fumigation has been used to control such insect damage. Additionally crop rotation practices can control *Diabrotica* populations. However, excessive use of insecticidal chemicals is not environmentally desirable and the practice of crop rotation may not be feasible for all growers. There is also increasing evidence that annual crop rotation may not provide complete protection from *Diabrotica* due to the increased prevalence of extended diapause in some *Diabrotica* populations. Therefore a means of protecting the plants from such insect damage by a mechanism other than chemical control or crop rotation would be very beneficial.

Proteases or peptidases, hereinafter referred to collectively as proteases, are enzymes which hydrolyze peptide bonds in proteins or peptides and demonstrate this activity at the ends of peptides (exopeptidases) or within the peptide chain (endopeptidases). The endopeptidases in particular cleave internal peptide bonds with different degrees of specificity for particular amino acyl residues. The protease enzymes are classified on the basis of their catalytic residues. Thus, with thiol proteases a cysteine sulfhydryl group participates directly in cleavage of the substrate peptide bond. Proteases are ubiquitous in nature and it is well documented that these enzymes are present in insect and other plant pests.

Peptides are generated from proteins in the gastrointestinal tract of organisms by the action of proteases during digestion. These organisms must rely on the activity of a battery of potent gut

proteases in order to obtain all amino acids. The best characterized insect proteases are serine proteases (Applebaum, S.W. (1985) Biochemistry of Digestion, in Comprehensive Insect Physiology. Biochemistry and Pharmacology, Kerkut, G.A. and Gilbert, L. eds. 4:279-311. Pergamon Press, Oxford). In contrast, less polyphagous insects may employ acid proteases (Pendola, S. and B. Greenburg, (1975) Ann. Ent. Soc. Am. 69:341-345) or thiol proteases (Murdock et al., (1987), Comp. Biochem. Physiol. 87B:783-787) as major contributors in protein digestion. The gut of *Diabrotica virgifera*, common name, western corn rootworm (wCRW) has been found to be particularly rich in thiol proteases (Murdock et al, supra).

Molecules which form complexes with proteases and inhibit their proteolytic activity are also widespread in nature and they are regulators of proteolytic activity. The presence of a peptide inhibitor of thiol proteases, known as cystatin, was first described in 1957. Since this initial discovery, many cystatin type inhibitors have been characterized, and these inhibitors are generally referred to as members of the cystatin superfamily. The three subgroups of the superfamily include: the cystatins, containing disulfide bond(s); the stefins, lacking cysteine residues; and the kininogens which are high molecular weight glycoproteins with disulfide bonds.

Protease inhibitors are among the defensive chemicals in plant tissue that are both developmentally regulated and induced in response to insect and/or pathogen attack. Suppression of serine protease inhibitor expression in transgenic tomato plants has been demonstrated to result in reduced tolerance to insect feeding.

Certain thiol protease inhibitors that are toxic to insects and in particular Coleoptera are well characterized. In three species of beetles studied by Murdock et al, (supra) it was found that much or most of the proteolytic activity in the midgut extracts was inhibited by E-64, [N-(L-3 transcarboxyoxiran-2-carbonyl)-L-leucyl]-amido(4-guanido)-butane, a specific, potent nonprotein inhibitor of thiol proteases.

The present invention is particularly concerned with the discovery of a novel peptide designated virgiferin as an effective thiol protease inhibitor. The amino acid sequence of virgiferin and

modified virgiferin distinguishes these peptides from thiol protease inhibitor members of the cystatin superfamily.

In accordance with the invention a thiol protease inhibitor, hereinafter referred to as virgiferin is provided which is about a 11.5 kDa peptide with an amino acid sequence beginning at amino acid position 1 and ending at amino acid position 83 of SEQ ID NO.: 1 and SEQ ID NO.: 2. Virgiferin is believed to be a novel and unique peptide. No peptide with the foregoing amino acid sequence and combination of properties is believed to have been described heretofore. SEQ ID NOS.: 1 and 2 also provide the amino acid sequence for the signal peptide extending from amino acid position -17 to position -1.

In accordance with another aspect of the invention are modified virgiferin peptides, said modified peptides are functionally equivalent to virgiferin and comprise peptides having substantial amino acid sequence similarity to virgiferin; truncations of virgiferin; and truncations of peptides having substantial amino acid sequence similarity to virgiferin.

Thus the invention also includes pure proteins which have a substantially similar activity/function to the protein depicted in Seq ID No. 2, and which have an amino acid sequence which is at least 50% similar to the sequence of the protein depicted in the said Seq ID. It is preferred that the degree of similarity is at least 70%, more preferred that the degree of similarity is at least 90% and still more preferred that the degree of similarity is at least 95%. In the context of the present invention, two amino acid sequences with at least 50% similarity to each have at least 50% identical or conservatively replaced amino acid residues in a like position when aligned optimally allowing for up to 5 gaps with the *proviso* that in respect of the gaps a total not more than 10 amino acid residues are affected. For the purpose of the present invention conservative replacements may be made between amino acids within the following groups:

- (i) Alanine, Serine and Threonine;
- (ii) Glutamic acid and Aspartic acid;
- (iii) Arginine and Lysine;
- (iv) Asparagine and Glutamine;

- (v) Isoleucine, Leucine, Valine and Methionine;
- (vi) Phenylalanine, Tyrosine and Tryptophan"

Where the term "truncation" is used in connection with a protein sequence, the term means a peptide comprised by the sequence depicted in SEQ ID No. 2 and having at least 10 amino acids. More preferably the peptide has at least 35 amino acids, still more preferably the peptide has at least 55 amino acids, and still more preferably the peptide has at least 70 amino acids.

In accordance with still another aspect of this invention are recombinant nucleic acid sequences which encode virgiferin, or modified virgiferin (defined hereinbelow), and nucleic acid sequences similar to those which encode the said virgiferin, or modified virgiferin. By similar is meant a sequence which is complementary to a test sequence which is capable of hybridizing to the inventive sequence. When the test and inventive sequences are double stranded the nucleic acid constituting the test sequence preferably has a  $T_M$  within  $10^\circ\text{C}$  of that of the inventive sequence. In the case that the test and inventive sequences are mixed together and denatured simultaneously, the  $T_M$  values of the sequences are preferably within  $5^\circ\text{C}$  of each other. More preferably the hybridization is performed under stringent conditions, with either the test or inventive DNA preferably being supported. Thus either a denatured test or inventive sequence is preferably first bound to a support and hybridization is effected for a specified period of time at a temperature of between  $50$  and  $60^\circ\text{C}$  in double strength citrate buffered saline containing  $0.1\%$  SDS followed by rinsing of the support at the same temperature but with a buffer having a reduced SC concentration. Depending upon the degree of stringency required, and thus the degree of similarity of the sequences, such reduced concentration buffers are typically single strength SC containing  $0.1\%$  SDS, half strength SC containing  $0.1\%$  SDS and one tenth strength SC containing  $0.1\%$  SDS. Sequences having the greatest degree of similarity are those the hybridization of which is least affected by washing in buffers of reduced concentration. It is most preferred that the test and inventive sequences are so similar that the hybridization between them is substantially unaffected by washing or incubation in one tenth strength sodium citrate buffer containing  $0.1\%$  SDS.

Further DNA constructs are provided comprising a promotor which functions in a host cell and a DNA sequence encoding one or more of said virgiferin peptides.

Also embodied by the invention is a method for controlling insect or nematode infestation which comprises producing genetically transformed plants which have decreased susceptibility to damage by insect or nematode pests comprising inserting into the genome of a plant cell a DNA construct of the invention, obtaining transformed plant cells and regenerating from the plant cells genetically transformed plants which have reduced susceptibility to damage by insect and nematode pests and particularly to damage by Coleoptera.

This method for controlling insect or nematode infestation may also include providing plant-colonizing microorganisms which have been transformed to express a gene for virgiferin or modified virgiferin and wherein said microorganisms are introduced to a plant or plant loci where control is desired and wherein said gene is expressed in an insecticidally or nematodically effective amount.

The invention also envisages propagating or regenerating a genetically transformed plant as defined above and using plant cells, tissue cuttings, or seed to produce progeny that also display reduced susceptibility to damage by insect pests and in particular to Coleopteran insects such as *Diabrotica*. The invention additionally covers the progeny produced by said propagation or regeneration.

Another embodiment of the invention includes the production of protein derived from the expression of virgiferin and modified virgiferin in host microbial cells, and in particular including *E.coli* and root-colonizing organisms.

Other objects and features of the invention will become apparent from the following detailed description. It should be understood that the detailed description and specific examples are by way of illustration only. In the examples:

Figure 1 describes the Laemmli gel of proteins extracted from wCRW pupae and subjected to

cm-papain chromatography;

Figure 2 describes an active site titration of wCRW gut proteases with the irreversible inhibitor E-64 and with two concentrations of MAN-virgiferin;

Figure 3 describes a plasmid map of the *E. coli* expression construct pEV4 encoding the virgiferin peptide;

Figure 4 describes a plasmid map of the plant expression construct pZO1731 encoding the virgiferin peptide;

Figure 5 describes a plasmid map of the plant expression construct pZO1732 encoding the virgiferin peptide which is fused to the PRMS (P14) leader peptide;

Figure 6 describes a plasmid map of the plant expression construct pZO1522 which is delivered simultaneously with each virgiferin encoding construct.

The novel thiol protease inhibitor, virgiferin, is defined by the amino acid sequence of SEQ ID NOS.: 1 and 2 starting at amino acid position 1 and ending at amino acid position 83. The position numbering is from the N-terminal serine (Ser) of virgiferin. Virgiferin is a potent inhibitor of *Diabrotica* digestive proteases. Pre-virgiferin is encoded as the precursor protein with a signal leader sequence of SEQ ID NOS.: 1 and 2 starting at amino acid position -17 to position -1. As used herein a "signal sequence", also referred to as a leader sequence, is about 11 to about 22 amino acids. These amino acids are cleaved during maturation of the protease inhibitor. Leader sequences are well known in the art and are specific N-terminal sequences known to efficiently direct the mature peptide to the endoplasmic reticulum, vacuole or extracellular space via translocation through the endoplasmic reticulum membrane and which are excised during translocation. Examples of leader sequences include not only the PRMS (P14) sequence used in pZO1732 herein but also those from endoproteinase B gene and the tobacco PRIG gene among others. As used herein the term protease includes proteinases, and exo and endo-peptidases and means enzymes which hydrolyse peptide bonds in proteins and



peptides. Further as used herein the term "protease inhibitor" refers to a molecule that reduces the level of proteolytic activity or esterase activity of a given protease enzyme.

The invention includes modified virgiferin peptides that may be isolated or constructed through standard techniques. Modified virgiferin peptides are functionally equivalent to virgiferin and have substantial amino acid sequence identity to virgiferin, or are truncations of virgiferin, or truncations of peptides with substantial amino acid sequence identity thereto (see above). A functionally equivalent peptide is one wherein one or more amino acids have been added, substituted or removed without substantially reducing the peptides' protease inhibitory activity as compared to virgiferin. One skilled in the art is aware that various amino acids can be replaced or deleted in a peptide and yet that peptide still maintains its function. In this regard, a modified functionally equivalent peptide may include peptides with an even higher level of insecticidal or nematodical activity than virgiferin. Such modified peptides are preferred to have at least 50% similarity to the amino acid sequence of virgiferin, preferably at least 70% similarity and more preferably at least 90%.

Modified virgiferin peptides and peptides of thiol proteases may be characterized by a molecular weight of about 10 - 15 kDa on an SDS PAGE with binding affinity to papain type proteases. Papain is well known to those skilled in the art as a thiol protease of defined three dimensional structure (Drenth, J et al., (1976) *Biochemistry* 15:3731-3738 ). A propeptide of a thiol protease is defined as the portion of the precursor thiol protease which is enzymatically cleaved from the catalytic protease peptide upon activation or maturation of the protease. In the case of papain type thiol proteases, propeptides occur amino-terminal to the enzymatic peptide. Additionally, modified virgiferin peptides may include peptides isolated from a Coleopteran insect and particularly a *Diabrotica*.

In this regard the invention also includes polymorphic clones of virgiferin. In particular at amino acid position 64 of SEQ ID NO.: 1, wherein proline or leucine may occur; and at amino acid position 73, where valine or leucine may occur. A summary of these polymorphisms is shown in Table 1.

The invention also includes peptides generated by truncation of virgiferin and virgiferin peptides modified either by genetic or chemical means while still maintaining the protease inhibitor function. Truncations refer to the elimination of amino acid sequences from the amino terminal and/or carboxy terminal ends of the peptide. Further the invention covers fusion proteins in which virgiferin sequences or modified virgiferin sequences are joined to the sequence of another protein or peptide either by chemical or genetic means. Fusion proteins may include either peptides joined in tandem or where the virgiferin sequence or modified virgiferin sequences are interrupted.

As noted one aspect of the invention is to produce virgiferin and modifications thereof through recombinant DNA techniques as well as to provide host cells that have been genetically transformed with nucleic acid sequences which code for virgiferin and modifications thereof. Host cells may include plant cells and bacterial cells.

DNA constructs suitable for transformation include at minimum the foreign gene including a promoter and coding sequence for virgiferin or modified virgiferin. The DNA sequences of the present invention can be obtained using various well known methods. The DNA may be synthesized by polymerase chain reaction, chemically synthesized, or from isolation of double stranded DNA sequence from genomic DNA or complementary DNA (cDNA) which contains the sequence. The cDNA sequence of virgiferin is described in SEQ ID No.: 1.

As is known in the art, the degeneracy of the genetic code allows for various nucleic acid sequences, DNA's and RNA's to encode the same protein, and in most cases an amino acid is encoded by two or more synonymous codons, for example the amino acid alanine is encoded by GCU, GCC and GCA. When cloning the peptides of the invention in another host organism, it may prove desirable to alter the nucleic acid codons such that those preferred by the host organism are employed, although no changes are made in the translation product. In this regard the invention also includes these synonymous codon sequences. The invention further includes not only preferential codon usage by a host organism but also changes in the polyadenylation sequence and intron splice consensus sequences.

The invention includes nucleic acid sequences which hybridize under stringent hybridization conditions with the cDNA virgiferin sequence or a portions thereof as disclosed herein. "Stringent hybridization conditions" are those in which hybridization is effected between 50 and 60°C in saline buffer solution (see above). The DNA to be used for hybridization may be prepared in a conventional manner and may be targeted to form an identifiable hybridization probe by procedures well known in the art. Such a probe may also be used to probe suitable and conventionally prepared clone banks to locate and isolate nucleotide alleles of virgiferin or to locate and isolate genes coding for virgiferin or modifications thereof in other Coleopteran insect pests. Particularly the nucleic acid segments or probes may be labeled chemically or radiochemically and used in Northern Hybridization, Southern Hybridization and colony or plaque hybridization to identify similar or identical segments from *Diabrotica spp.*

Vectors comprising the nucleic acid sequences described above represent another embodiment of the invention. Expression vectors are typically plasmids. Plasmids in general are circular double stranded DNA loops including a promotor operably linked to the DNA sequence or sequences encoding a protein of interest, transcription termination sequences and the remaining vector with 3' and 5' elements. One skilled in the art is aware of many types of vectors including virus vectors, baculovirus; phage vectors; Agrobacterium-Ti plasmid; binary vectors and other vectors suitable for plant transformation. The insertion of the DNA sequences into microbial and plant host cells according to the invention occurs according to techniques known in the art.

Promoters refer to nucleotide sequences at the 5'-end of a structural gene which direct the initiation of transcription and include all the regulatory regions required for transcription including the region coding for the leader sequence of mRNA. A number of promoters which are active in microbial and plant cells have been described in the literature. Suitable plant promoters include: promoters of the genes encoding chlorophyll a/b binding proteins, nopaline synthase (NOS), octopine synthase (OCS), cauliflower mosaic virus (CaMV) 19S and 35S, ribulose biphosphate carboxylase (RUBISCO), and heat shock *Brassica* promotor (HSP 80). Suitable promoters used in the transformation of *E. coli* include, but are not limited to  $P_{Tac}$ ,  $\lambda_{pr}$  and  $T_7$  which are available commercially. The promoters used in the present invention may be

modified to affect control characteristics and further may be a composite of segments derived from more than one source, naturally occurring or synthetic. Termination sequences refer to a nucleotide sequence at the end of a transcriptional unit that signals termination of transcription. Terminators are 3'-non-translated DNA sequences that contain a polyadenylated signal. Examples of terminators are known and described in the literature and include but are not limited to nos (nopaline synthase terminator), the 35S terminator of CaMV and the zein terminator.

Vectors of the invention may also possess other DNA sequences known in the art, including but not limited to: a) stability sequences; b) one or more marker sequences, for example, antibiotic or herbicide resistance markers including cat (chloramphenicol acetyl transferase), npt II (neomycin phosphotransferase II), PAT (phosphinothricin acetyltransferase the expression of which confers resistance to the herbicide Basta); EPSPS ( 5-enolpyruvyl- shikimate 3-phosphate synthase, the enzyme inhibited by glyphosate, the active ingredient in the herbicide Roundup) and bxn (bromoxynil-specific nitrilase); c) other insecticidal protein genes such as delta endotoxin sequences for a *Bacillus thuringiensis*; d) intron sequences; and e) enhancer or other elements necessary to increase or decrease levels of expression obtained in particular parts of plants under certain conditions. These examples are not intended to limit the invention in any manner.

A wide variety of techniques are available and known in the art for carrying out the actual transformation and transfection. These include but are not limited to, direct transfer of DNA into whole cells, tissues or protoplasts, optionally assisted by chemical or physical agents to increase cell permeability to DNA, for example, treatment with polyethylene glycol, and dextran sulfate; electroporation, heat shock and ballistic implantation of DNA coated particles. Transformation is also mediated by *Agrobacterium* strains, notably *A. tumefaciens*, and also by various genetically engineered transformation plasmids which include portions of the T-DNA of the tumor inducing plasmids of *Agrobacteria*. Other means for effecting entry of DNA into cells include the use of viral vectors, agroinfection, binary vectors and cotransformation. (See Jensen et al., (1993), Techniques for Gene Transfer, pp 125 -146 in Transgenic Plants, vol. 1, eds. King and Wu, Academic Press). Transformed or transfected bacterial cells are included

in the present invention, for example *E. coli* and *Bacillus thuringiensis* (*B.t.*). Considerable experience in biotechnology has already been achieved and a wide variety of suitable operatively functional plasmids and transfer expression vectors systems are known and available for these organisms.

As used herein the term "genetic transformation" means the stable integration of a foreign gene into the genome of a plant regenerated from nucleic acid treated plant protoplasts, cells or tissue. "Transgenic plants" as used herein refers to plants carrying the stably integrated foreign gene. A "foreign gene" is a term used in the art to denote a gene or group of genes which has been transferred to a host cell or host plant from a source other than the host cell or host plant. The terms "transgene" or "exogenous gene" have the same meaning as the term foreign gene as used in this application.

Virtually all plants of agronomic or horticultural value are known to be both transformable and regenerable. The techniques vary in individual detail from species to species as is well recognized by one skilled in the art. Means of regenerating plants are well documented in the literature. For a review on plant transformation and regeneration, see Ritchie and Hodges, pps. 147-178, in Kung and Wu, *Transgenic Plants*, vol. 1, 1993, Academic Press. Plant tissue includes differentiated and undifferentiated tissue and cells of plants including but not limited to roots, shoots, leaves, pollen, embryos, seed and various forms of aggregations of plant cells including callus.

The present invention involves the preparation of recombinant DNA molecules that function in host cells and particularly in plant cells to produce decreased susceptibility to damage by plant insect pests and other plant pests that utilize cysteine digestive proteases. Such insect pests include in particular insects of the families *Teuebrionidae*, *Curculionidae*, *Bruchidae* and *Chrysomelidae*. Mention is made of insects of the genus *Diabrotica*. Exemplary of such plant insect pests are the species *adelpha*, *balteata*, *barberi*, *speciosa*, *undecimpunctata*, *howardi*, *virgifera*, and *viridula* commonly referred to as banded cucumber beetle, northern corn rootworm, spotted cucumber beetle, southern corn rootworm, western corn rootworm and Mexican corn rootworm. Other specific insects include Mexican bean beetle, Colorado potato

beetle, wireworms, grubs, bollweevils, maize, grainery and rice weevils and Bruchids.

Decreased susceptibility to damage by plant pests may be provided by expressing one of the virgiferin peptides of the invention in a root-colonizing organism wherein the organism is applied to the plant or to a loci around the plant and as the pest feeds on the plant it ingests an insecticidally effective amount of the virgiferin peptide produced by the organism. Root-colonizing organisms may be bacterial or fungal and further may be surface colonizers or endophytes.

Plants which may be made less susceptible to damage by insect or nematode pests by the practice of the present invention include but are not limited to those plants that are members of Gramineae, Cucurbitaceae, Leguminosae, Solanaceae, Compositeae, Cyperaceae, Convolvulaceae. Specific mention is also made of the following plants: corn (maize), sweet corn, squash, melon, cucumber, sugarbeet, sunflower, rice, cotton, canola, sweet potatoes and legumes including peanuts, common bean, cowpea, soybean, alfalfa and broad bean.

One further aspect of this invention comprises an insecticidal composition comprising as its active ingredient an insecticidally effective amount of a peptide as claimed in the present invention and in particular virgiferin. The term "insecticidally effective amount" or "nematodical effective amount" means an amount sufficient to achieve control of an insect or nematode plant pest by reducing the number of insects or nematodes through mortality, reduced growth (stunting), reduced reproductive efficiency and the like. Virgiferin or a modification thereof may be formulated in a number of ways well known to those skilled in the art. Formulations used in spray form such as water dispersible concentrates or wettable powders, may contain surfactants such as wetting agents or dispersing agents. Further active ingredients, for example pesticides including fungicides, insecticides and herbicides may be applied together with the peptides of the invention in order to increase their spectrum of activity or agricultural utility. The formulations may be applied at a locus where control is desired. The amount of active ingredient will vary depending on many factors, including the nature of the particular formulation.

The following examples are provided to illustrate the invention and are not intended to limit its scope. The examples utilize many techniques well known in the field of molecular biology, particularly with respect to manipulation of recombinant DNA in plant tissue with the concomitant transformation of plant tissue. The techniques are described in further detail by the cited references which are hereby incorporated by reference.

Example 1: Isolation and Purification of Virgiferin from wCRW pupae.

Virgiferin is purified from homogenates of wCRW pupae based on cm-papain-agarose chromatography. (Barrett, A. J. (1981) Methods in Enzymology 80: 771 - 778).

Preparation of cm-papain-agarose. Papain (Boehringer-Mannheim) is inactivated by carboxymethylation as described in Barrett, supra, and then dialyzed overnight in 50 mM MOPS (pH 7.5) -10 mM iodoacetic acid. Cm-papain is coupled to Affi-Gel 15 agarose (BioRad) in 50 mM MOPS (pH 7.5) and washed successively with 50 ml PBS, 50 ml 50 mM sodium carbonate buffer (pH 10.5) and 50 ml PBS prior to storage in PBS containing sodium azide.

Pupae homogenate. Approximately 50 wCRW pupae (French Agricultural Research) are washed with PBS to remove field soil. The pupae are homogenized in 15 ml PBS and the homogenate is cleared of debris by centrifugation. The supernatant is heated to 56°C for 5 minutes then centrifuged. The supernatant is mixed with 1.0 ml of cm-papain-agarose overnight at 4°C. The slurry is poured into a 10 ml disposable column and washed with PBS until detectable protein is not evident in the eluant. The column is washed with 50 mM sodium carbonate buffer (pH 10.5). One ml fractions are collected and evaluated by visualization on Laemmli electrophoresis through 4 - 20 % acrylamide gels for the presence of distinct peptides. A 11.5 kDa peptide is present as shown in Figure 1 and the peptide is designated virgiferin.

Example 2: Determination of the virgiferin N-terminal amino acid sequence.

Cm-papain fractions containing relatively high concentrations of the virgiferin are pooled and lyophilized. Virgiferin is resolved from minor contaminants by SDS-PAGE as described previously. Using the method developed by Matsudaira (Matsudaira, P. (1987) J Biol Chem 262: 10035 - 10038), virgiferin is electrophoretically transferred to a polyvinylidene difluoride

membrane (Millipore) adjustments are made for transfer of low molecular weight proteins as follows. The acrylamide gel is soaked in transfer buffer (40% methanol, 10 mM TRIS (pH 8.8)) for one hour. Proteins are electrophoretically transferred overnight at 30 mA in transfer buffer to a PVDF membrane. An intense band of protein is visualized upon staining the membrane. The band is excised and subjected to automated protein sequence analysis. The sequence is determined by automated Edmond degradation and presented SEQ ID NO.: 3. In addition to the amino acid residues listed in SEQ ID NO.: 3, other residues are present in lesser amounts and these include in addition to Ser (amino acid residue 1), Asn and Ala; in addition to Gln (amino acid residue 2), Ile and Pro; in addition to Thr (amino acid residue 3), Ala and Leu; in addition to Glu (amino acid residue 5), Gly; in addition to Gly (amino acid residue 8) Asp; and in addition to Ala (amino acid residue 17), Arg.

### Example 3: cDNA Synthesis and Cloning

PCR techniques are used to produce cDNA clones encoding virgiferin. Unless specified the starting templates for the PCR reactions are RNA purified from wCRW tissues as follows:

RNA isolation and polyA+ selection: Total RNA for PCR and Northern blot analysis is isolated from WCR third-instar gut and pupae. Tissue is homogenized in RNazol (Tel-Test) following supplier's protocol. PolyA+ RNA is purified from total RNA by Poly(A)Quik oligo-dT cellulose (Stratagene).

Polymerase Chain Reactions: Polymerase chain reactions (PCR Protocols) are performed on a Perkin-Elmer Cetus 9600 DNA Thermal Cycler. Reaction conditions are as follows: 2U Taq DNA polymerase (Pharmacia) in enzyme's buffer, 0.2 uM dNTP mix, 0.4 uM each of sense and antisense primer, and DNA template in water to a final volume of 50 ul. Template in water is denatured at 94C/1 min prior to the addition of the reaction mix at 82°C. The reaction cycle is 94C/15 sec, 55C/30 sec, 72C/45 sec for 30 cycles, followed by a 72C/3 min extension.

Template DNA Synthesis: Template DNA for PCR is first-strand cDNA synthesized using the First-strand cDNA Synthesis Kit (Pharmacia). PolyA+ RNA (wCRW gut and pupal) are primed with the Not I-d(T)18 Bifunctional Primer from the kit. The sense primer for the initial PCR is a degenerate oligonucleotide, P K Y:



5'GA(A/G)GA(A/G)GC(A/C/G/T)G(A/G)(A/C/G/T)CC(A/C/G/T)AA(A/G)TA(T/C) - AA(A/G)AC 3', SEQ. ID. NO.: 4, encoding the N-terminal amino acid sequence of the 11.5 kDa wCRW protein as determined above.

The antisense primer, TLESS: 5' AACTGGAAGAATTCGCGGCCGCAGGAA 3' (SEQ. ID. NO.: 5) is based on a portion of the Not I-d(T)18 Bifunctional Primer.

A 600 base pair PCR product is obtained.

Purification and Cloning of PCR Products: PCR products are gel-purified in a 1.2% agarose gel in TBE. DNA is eluted from the excised gel fragment by electroelution. Gel-purified PCR products are ligated into the T-vector pT7Blue(R)(Novagen) using standard methods (Maniatis, supra).

DNA Sequencing: Recombinant clones are selected using PCR techniques and cultured in 3 ml volumes. DNA is prepared from these cultures and are sequenced using standard methods. The sense primer for sequencing templates cloned in the plant expression vectors is IVS6END: 5' CTGAATTTGTGAACCCAA 3' (SEQ ID NO.: 6) based on the 3' end of intron 6, and the antisense primer is OCS rev: 5' GAATTGAAAGCAAATATCATGCG 3' (SEQ ID NO.: 7) based on the 5' end of the complete OCS terminator. These two primers flank the cloning sites of the plant expression vectors, pZO1731 and pZO1732 (Figure 4 and 5 respectively). The construct pZO1732, DNA encoding the PRMS(P14) signal sequence is fused in frame with virgiferin as described by amino acid position 1 to 83 in SEQ ID NO.: 1. (Casacuberta, et al., (1991) Plant Mol. Biol. 16:527-536).

Cloning the 5'-end of the Virgiferin cDNA: Two nested antisense primers are designed for cloning the 5' end of the gene from the sequence data of the initial PCR-amplified DNA: VTW rev: 5' CATGGACCATGTTACTTCACCAGCTTCG 3' (SEQ. ID. NO.: 8) and DQ rev: 5' GTTGTGTGCTTCGATCTGGTCG 3' (SEQ. ID. NO.: 9). The 5' ends are cloned using the 5'RACE System (Gibco BRL) following manufacturer's protocol. The complete DNA sequence and deduced amino acid sequence of the propeptide (pre-virgiferin) is shown in SEQ

ID NO.: 1 and includes the 17 amino acid leader peptide encoded in the 5' DNA.

Example 5. Expression of Virgiferin in *E. coli*.

For construction of expression vectors, PCR primers are synthesized that introduce a methionine start codon and appropriate restriction sites. The sense primer SQT: 5' GCGCGTCGACCATGGCGAATTCACAACTGCTGAAGAAGC 3' (SEQ. ID. No.: 10) introduces NcoI and EcoRI sites and the residues MAN just N-terminal to the first amino acid of the virgiferin peptide (starting at the Ser residue, amino acid position 1 in SEQ ID NO.:1). The MAN residues represent the amino acids methionine, alanine and asparagine respectively. The antisense primer; GVH rev: 5' CCGGCTGCAGAAGCTTTTAATGTACTCCAGTAGCACC 3' (SEQ. ID. NO.: 11) introduces HindIII and PstI sites after the stop codon. Both primers include four nucleotide "clamps" at their 5' end for facilitating restriction digests.

The DNA sequence encoding virgiferin is amplified using PCR technology from RNA isolated from wCRW gut as described previously, using the SQT and GVH rev primers. These primers introduce a methionine start codon and appropriate restriction sites. The resultant PCR products are purified and ligated into pT7Blue(R), screened and sequenced as before.

The DNA sequence reveals three polymorphisms in the population of clones: at amino acid position 64, proline or leucine; position 73, valine or leucine; and position 75, proline or arginine. For position 73 and 75, only one clone, V7-3, has leucine and arginine, respectively. The valine to leucine transition requires two base changes and the proline to leucine transition requires two base changes (CCA to TTA). Two polymorphic clones, V7-3 and V7-9 have been characterized and further developed in bacterial and plant expression constructs. A summary of these polymorphisms is shown in Table 1.

**TABLE 1**

Clone	ResidueExpression Vectors Based on Clone
-----	-----
	647375 BacterialPlant
V7-3	ProLeu ArgBN8VGP322(pZO1730)
V7-9	ProValProPEV4VNP31(pZO1731)
	VRE92(pZO1732)

The pT7Blue(R) construct encoding virgiferin with N-terminal MAN residues, described above, is digested with restriction enzymes EcoRI and HindIII. The DNA fragment encoding virgiferin is purified and ligated into the *E. coli* expression vector pET-20b(+)(Novagen), thus fusing the DNA sequence encoding MAN-virgiferin in frame with the DNA sequence encoding the *E. coli* pelB leader sequence. Transformed cultures are screened using PCR methods as described previously. DNA sequence analysis of PCR-positive clones confirmed in-frame fusions of the pelB leader with MAN-virgiferin. This construct is referred to as plasmid pEV4 and described in Figure 3.

pEV4 is used to transform *E. coli* strain BL21(DE3). Freshly transformed colonies are transferred to liquid media and cultured as described Dubendorf, J.W. and Studier, F.W., (1991), J.Mol.Biol., 219:45-59 and Studier F.W. (1991) J. Mol. Biol. 219:37-44. Aliquots of the culture are removed at the time of induction and at intervals of 30 minutes, 60 minutes, two hours and four hours following induction and analyzed by SDS-PAGE. Abundant expression of a 10 - 12 kDa protein is apparent in samples taken two hours or later following induction.

#### Example 6. Purification of Recombinant Virgiferin.

For preparative purposes, 500 ml bacterial cultures are incubated 4 hours at 37 C following induction for optimal yield. The cultures are harvested and hypotonically shocked to lyse the outer cell wall. See Lin, K. and Cheng, S. (1991) BioTechniques 11: 748 - 752. The spheroplasts are separated from the periplasmic protein-enriched fraction by centrifugation. The periplasmic fraction contains nearly pure MAN-virgiferin which is stable at 4°C for up to four weeks. In some cases this crude preparation of MAN-virgiferin is tested for proteolytic

inhibitory activity.

For chromatographic purification of MAN-virgiferin 1 M Tris-HCl (8.0) is added to the periplasmic fraction to make a final concentration of 50 mM Tris-HCl (8.0). The supernatant is applied to a DEAE Sepharose Fast Flow (Pharmacia) column in 50mM Tris-HCl (8.0) at a flow rate of 3 ml/min. Proteins are eluted with a 0mM - 350mM NaCl linear gradient in the same buffer. Column fractions are analyzed by SDS-PAGE. MAN-virgiferin is eluted from the DEAE column in a single peak at a relatively high NaCl concentration. The DEAE column fractions containing virgiferin are pooled and dialyzed in Spectrapor #1 (MW cutoff = 6-8 kDa) tubing against 50mM Phosphate, 150mM NaCl (pH 7.0) for 2 hours. The protein is concentrated by ultrafiltration and applied to a Superose 12 (Pharmacia) gel filtration column. The Superose column is developed with the 50mM phosphate, 150mM NaCl buffer at a flow rate of 0.3 ml/min. SDS-PAGE analysis of the Superose fractions verifies that MAN-virgiferin has been separated from a minor contaminant. The protein concentration of Superose fractions containing pure MAN-virgiferin is determined by Bradford assay and the fractions are tested directly for inhibitory activity against wCRW gut proteases.

Antibodies have been produced from the *E. Coli* virgiferin protein and have been shown to be of high quality and effective for Western blot analysis.

#### Example 7. Proteolytic Inhibition Activity of Virgiferin.

Midguts, free of fat bodies are dissected from wCRW third instar larvae (French Agricultural Research) by evisceration. The wCRW midguts are then suspended in 100 mM MES (pH 6.3) (1 gut/ $\mu$ l). Following a 60 minute incubation on ice, the tissue is pelleted by centrifugation (13 K rpm, 8 min) and the supernatant is collected for analysis. This fraction contains potent thiol protease activity(s) effective in a broad range of pH 5.0 to pH 8.0.

Enzymatic activity of all preparations is measured against azocasein as described by Mason, R.W., Gal, S. and Gottsman, M.M. (1987) *Biochem J.* 248: 449 - 454. Active site titrations for all enzymatic preparations are performed using azocasein as the proteolytic substrate. All assays are performed at 37°C for 60 minutes. The titration curves are normalized for full activity of

each enzyme preparation in the absence of thiol protease inhibitors. Enzyme inhibition is reported as a proportion of full proteolytic activity under the same conditions in the presence of designated quantities of the irreversible thiol protease inhibitor E-64 or MAN-virgiferin. A typical active site titration of wCRW gut proteases is shown in Figure 2. Repeatedly the titration curves obtained with MAN-virgiferin suggest potent inhibition of the wCRW gut proteases.

Example 8: Transformation and Regeneration of Maize Tissue

Corn ears, Hi Type II, 9 to 12 days after pollination are harvested for embryo isolation, sterilized and rinsed in sterile water. Embryos are excised aseptically from the kernels (Rhodes, C.A. et al., (1988) Bio/Tech. 6:56-60). Ten to 50 embryos are transferred to a solid N6ap1D medium, and stored in the dark at 27°C for 0 to 5 days.

Embryos are transferred to petri dishes containing solid N6ap1D medium supplemented with 0.5M mannitol. Approximately 4 hours after transfer the embryos are subjected to particle bombardment DNA delivery or "shot" with the Biolistic PDS-1000/He Particle Delivery System using a standard Sanford protocol except that the ethanol (100%) is added to the particles so that the final volume of the suspension is 900µL.

The virgiferin encoding plasmids used in the transformation of maize tissue are described in Figures 4 and 5 and correspond to pZO1731 and pZO1732. pZO1731 includes the virgiferin encoding sequence, a 35s CaMV promoter, intron VI, and ocs termination sequences. pZO1732 includes the signal leader PRMS in addition to the sequence encoding virgiferin. Each virgiferin encoding plasmid is delivered simultaneously with plasmid pZO1522 depicted in Figure 6. pZO1522 encodes the *Streptomyces hygroscopicus* phosphinothricin acetyltransferase (PAT) gene, the expression of which confers Basta resistance to transformed tissue. (Thompson C. J. et al., (1987) EMBO J. 6:2519-2523). Basta, common name glufosinate-ammonium, is a chemical herbicide containing phosphinothricin which inhibits glutamine synthase. The practice of co-shooting or mixing plasmid DNA's for particle delivery is known and described in the art (Gordon-Kamm et al., 1990, Plant Cell 2:603-681). The final volume of the suspension containing DNA-coated tungsten particles and ethanol is 900 µL

prior to delivery to the embryos.

Following particle delivery the embryos are kept on N6ap1D medium supplemented with 0.5M mannitol for approximately 16 hours then transferred to fresh N6ap1D medium without selective agents for 0 to 5 days. Finally, the embryos are transferred to solid N61D medium supplemented with 13 to 24mg/L of Basta. Transfers are made to fresh selective media every 2 weeks or as needed.

Sixty days after transformation, callus and embryo tissue which continues to grow on N61D medium containing Basta are transferred to fresh medium containing 24mg/L Basta. Samples of at least 50ug of these transformant tissues are removed and analyzed for the presence of the DNA originally delivered by particle bombardment using PCR techniques.

To regenerate plants the callus is transferred to petri dishes containing solid MSap medium without growth hormones and stored in the dark at 27 C for 15 days. Plant shoots formed under these conditions are transferred to fresh medium in boxes and stored at 27 C under a cycle of 18 hours/24 hours of bright light. Ten days after transfer to boxes the developing plants that are greater than 3 inches in height and possess a root system are transferred to soil and stored in a growth chamber for 48 hours total. The relative humidity is adjusted to 50% and the plants maintained in a growth chamber for 1 to 2 weeks then transplanted into 5 gallon pots and maintained in a greenhouse with under a cycle of 18 hours/24 hours of bright light. Fertile, transformed plants are pollinated with control pollen and seed is collected.

Stability and segregation of the virgiferin gene is determined by analysis of germinated seedlings. Some of these results are indicated below. Each independently transformed line is shown to have the virgiferin gene by PCR analysis.

**TABLE 2**

	# of Stable Inherited Events	# of Events Showing Protein Expression
pZ01731	187	
pZ01732	153	

Using the antibodies produced against the virgiferin protein in Example 5, transgenic plants are screened for the production of virgiferin. Protein expression is determined by Western blot analyses according to standard techniques (Antibodies, A Laboratory Manual, 1988, Ed. Harlow and David Lar, Cold Spring Harbor Laboratory). Western analyses also indicates that the protein is produced in both leaves and roots of transformed plants with both gene constructs.

Example 9: Insect Susceptibility Tests

wCRW eggs (200 - 500) are disinfected and placed in a sterile environment to hatch. Larvae which hatch in 1 to 3 days are transferred to maize callus. Seven days following initial exposure of the larvae to the callus insecticidal activity is evaluated. The number of dead insects as well as the size of the surviving insects is recorded.

Example 10: Artificial Diet Bioassays

Neonate wCRW larvae are allowed to feed on BioServe, a standard artificial diet used for southern corn rootworm. The diet is supplemented with the treatments indicated below in Table 3. Two or 3 larvae are placed in treatment wells with 9 replicate wells used per treatment. The larvae are allowed to feed on the diet and mortality is evaluated at day 6. The results of the assay are given below.

**TABLE 3**

<u>Treatment</u>	<u>% wCRW mortality <math>\pm</math> SE</u>
virgiferin (4%)*	100 $\pm$ 0
E64 (0.01%)	80 $\pm$ 8
Standard Diet	15 $\pm$ 8
Standard Diet (10%)+ Casein	29 $\pm$ 15

\*Virgiferin was obtained from the production of the protein in *E. Coli*. Estimated dose of virgiferin protein (w/v).

To confirm that the virgiferin effects are due to direct ingestion of the protein and not due to an antifeedant response. Food coloring, commercially available green dye, was added to the diet. Results indicate that the dye is readily taken up and is observed through the insects cuticle upon ingestion.

Example 11: Effect of Virgiferin against *Callosobruchus maculatus* (cowpea weevil)

A dose response analysis of the effects of virgiferin was performed on *Callosobruchus maculatus*, the seed feeding cowpea weevil. Using an artificial seed bioassay system and an ultrasonic detector as disclosed in Shade, R.E. et al. (1986), Environ. Entomol. 15: 1286-1291, disruption of the cowpea weevil's normal feeding pattern is determined by differences in mean feeding events. The results are indicated below in Table 4.

**TABLE 4 Effect of virgiferin<sup>a</sup> on feeding activity of cowpea weevil**

<u>Treatment</u>	<u>Mean<sup>b</sup></u>	<u>Dose % w/w<sup>c</sup></u>	<u>Count<sup>d</sup></u>	<u>Feeding Events</u>	<u>LSC<sup>e</sup> (0.05)</u>
Virgiferin	0.4	4	43	a	
Virgiferin	0.2	5	70	a,b	
Virgiferin	0.1	6	60	a	
Virgiferin	0.05	5	61	a	
Virgiferin	0.025	5	85	a,b	
Virgiferin	0.031	5	107	b	
Control	0.0	6	105	b	



- a. Recombinant virgiferin is produced in *E. coli* as described in Example 5.
- b. Mean number of feeding events detected in 30 sec samples.
- c. Fishers Protected LSD - means followed by the same letter are not significantly different.
- d. Count = Number of insects tested.
- e. Estimated dose of virgiferin protein in diet.

The results indicate that there was a reduction in the mean number of feeding events observed and this initial analysis suggests that virgiferin is active against cowpea weevil.

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

## (i) APPLICANT:

(A) NAME: SANDOZ LTD  
(B) STREET: Lichtstrasse 35  
(C) CITY: Basel  
(D) STATE: BS  
(E) COUNTRY: Switzerland  
(F) POSTAL CODE (ZIP): CH-4002  
(G) TELEPHONE: 061-324-1111  
(H) TELEFAX: 061-322-7532  
(I) TELEX: 965-05055

(A) NAME: Sandoz Patent GMBH  
(B) STREET: Humboldtstrasse 3  
(C) CITY: Loerrach  
(E) COUNTRY: Germany  
(F) POSTAL CODE (ZIP): D-7850

(A) NAME: Sandoz Erfindungen Verwaltungsgesellschaft  
MBH  
(B) STREET: Brunner Strasse  
(C) CITY: Vienna  
(E) COUNTRY: Austria  
(F) POSTAL CODE (ZIP): A-1230

(ii) TITLE OF INVENTION: Novel thiol protease inhibitor

(iii) NUMBER OF SEQUENCES: 11

## (iv) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

## (2) INFORMATION FOR SEQ ID NO: 1:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 364 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(iii) HYPOTHETICAL: NO

## (ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION: 63..364

## (ix) FEATURE:

(A) NAME/KEY: sig\_peptide  
(B) LOCATION: 63..113  
(D) OTHER INFORMATION: /codon\_start= 63

## (ix) FEATURE:

(A) NAME/KEY: mat\_peptide  
(B) LOCATION: 114..364  
(D) OTHER INFORMATION: /codon\_start= 114

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

CATCAGTTAA TCAACTGAAT TACTTGGTTA ATAAACTTCC AGTTCCTCTT CCACAATCAA 60

AA ATG TTT TGT AAA GTA TTT GTA CTT GCT CTC ATT GTA GCT GTT GCT 107  
 Met Phe Cys Lys Val Phe Val Leu Ala Leu Ile Val Ala Val Ala  
 -17 -15 -10 -5

GTT GCA TCA CAA ACT GCT GAA GAA GCA TGG CCA AAA TAC AAG ACT GAT 155  
 Val Ala Ser Gln Thr Ala Glu Glu Ala Trp Pro Lys Tyr Lys Thr Asp  
 1 5 10

TAC AAT AGG AAC TAC GAC GCT CAG GAA GAT GCT ACA AGA TTC GCC ATT 203  
 Tyr Asn Arg Asn Tyr Asp Ala Gln Glu Asp Ala Thr Arg Phe Ala Ile  
 15 20 25 30

TTC AAG ACT AAC TAC GAC CAG ATC GAA GCA CAC AAC AAA AAG TTC GAA 251  
 Phe Lys Thr Asn Tyr Asp Gln Ile Glu Ala His Asn Lys Lys Phe Glu  
 35 40 45

GCT GGT GAA GTA ACA TGG TCC ATG GGA TTG AAC CAA TTC GCT GAC AGG 299  
 Ala Gly Glu Val Thr Trp Ser Met Gly Leu Asn Gln Phe Ala Asp Arg  
 50 55 60

ACC TTA GAA GAA CTA AAA CAT CTT CAT GGA GTT AGA CCT CCA GTA GGT 347  
 Thr Leu Glu Glu Leu Lys His Leu His Gly Val Arg Pro Pro Val Gly  
 65 70 75

GCT ACT GGA GTA CAT TA 364  
 Ala Thr Gly Val His  
 80

## (2) INFORMATION FOR SEQ ID NO: 2:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 100 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Phe Cys Lys Val Phe Val Leu Ala Leu Ile Val Ala Val Ala Val  
 -17 -15 -10 -5

Ala Ser Gln Thr Ala Glu Glu Ala Trp Pro Lys Tyr Lys Thr Asp Tyr  
 1 5 10 15

Asn Arg Asn Tyr Asp Ala Gln Glu Asp Ala Thr Arg Phe Ala Ile Phe  
 20 25 30

Lys Thr Asn Tyr Asp Gln Ile Glu Ala His Asn Lys Lys Phe Glu Ala  
 35 40 45

Gly Glu Val Thr Trp Ser Met Gly Leu Asn Gln Phe Ala Asp Arg Thr  
 50 55 60

Leu Glu Glu Leu Lys His Leu His Gly Val Arg Pro Pro Val Gly Ala  
 65 70 75

Thr Gly Val His  
 80

## (2) INFORMATION FOR SEQ ID NO: 3:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Ser Gln Thr Ala Glu Glu Ala Gly Pro Lys Tyr Lys Thr Asp Tyr Asn  
 1 5 10 15

Ala

(2) INFORMATION FOR SEQ ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 26 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

GARGARGCNG RNCNAARTA WAARAC

26

(2) INFORMATION FOR SEQ ID NO: 5:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 27 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

AACTGGAAGA ATTCGCGGCC GCAGGAA

27

(2) INFORMATION FOR SEQ ID NO: 6:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

CTGAATTTGT GAACCCAA

18

(2) INFORMATION FOR SEQ ID NO: 7:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 23 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(iii) HYPOTHETICAL: NO

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:  
GAATTGAAAG CAAATATCAT GCG 23
- (2) INFORMATION FOR SEQ ID NO: 8:
- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 28 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- (iii) HYPOTHETICAL: NO
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
CATGGACCAT GTTACTTCAC CAGCTTCG 28
- (2) INFORMATION FOR SEQ ID NO: 9:
- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 22 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- (iii) HYPOTHETICAL: NO
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:  
GTTGTGTGCT TCGATCTGGT CG 22
- (2) INFORMATION FOR SEQ ID NO: 10:
- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 40 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- (iii) HYPOTHETICAL: NO
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:  
GCGCGTCGAC CATGGCGAAT TCACAACTG CTGAAGAAGC 40
- (2) INFORMATION FOR SEQ ID NO: 11:
- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 37 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- (iii) HYPOTHETICAL: NO
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:  
CCGGCTGCAG AAGCTTTTAA TGTACTCCAG TAGCACC 37

### CLAIMS

1. An inhibitor peptide - designated Virgiferin - of thiol proteases, said peptide having the amino acid sequence extending from amino acid position 1 to 83 of SEQ ID No.: 1 or a modified virgiferin peptide wherein said modified peptide comprises a peptide having substantial amino acid identity to virgiferin; truncations of virgiferin; or truncations of the peptide having substantial amino acid identity to virgiferin wherein said modified virgiferin is functionally equivalent to said virgiferin protease inhibitor.
2. A pre-virgiferin peptide characterized by the entire amino acid sequence given in SEQ ID NO.: 1 or peptides having substantial amino acid sequence identity thereto.
3. A homogenous protease inhibitor peptide characterized by a molecular weight of about 10 - 15 kDa by SDS PAGE with binding affinity to papain type proteases.
4. A recombinant DNA molecule comprising a nucleic acid sequence encoding the peptide according to any one of claims 1-3.
5. A recombinant molecule according to claim 4 characterized in that it is a vector.
6. Protein derived from the expression of the recombinant molecule according to claim 4.
7. A nucleic acid sequence comprising a sequence encoding the peptide of claim 1 or a nucleic acid sequence complementary to one which hybridizes under stringent hybridizing conditions to a nucleic acid sequence encoding the peptide of claim 1.
8. A transgenic plant comprising a foreign gene which encodes a peptide according to any one of claims 1-3.
9. A process for producing genetically transformed plants which have decreased susceptibility to damage by insect pests comprising

- a) inserting into the genome of a plant cell a foreign gene which further comprises the recombinant molecule according to claim 5;
- b) obtaining transformed cells; and
- c) regenerating from the transformed cells genetically transformed plants which have reduced susceptibility to damage by said pests.

10. A method of combatting insect pests which comprises exposing them to an insecticidally effective amount of a peptide according to any one of claims 1 to 3, wherein said peptide is expressed in a plant or plant colonizing micro-organism as a result of genetic transformation.

11. An insecticidal composition comprising an insecticidally effective amount of a peptide according to claim 1 and an agriculturally acceptable carrier thereof.

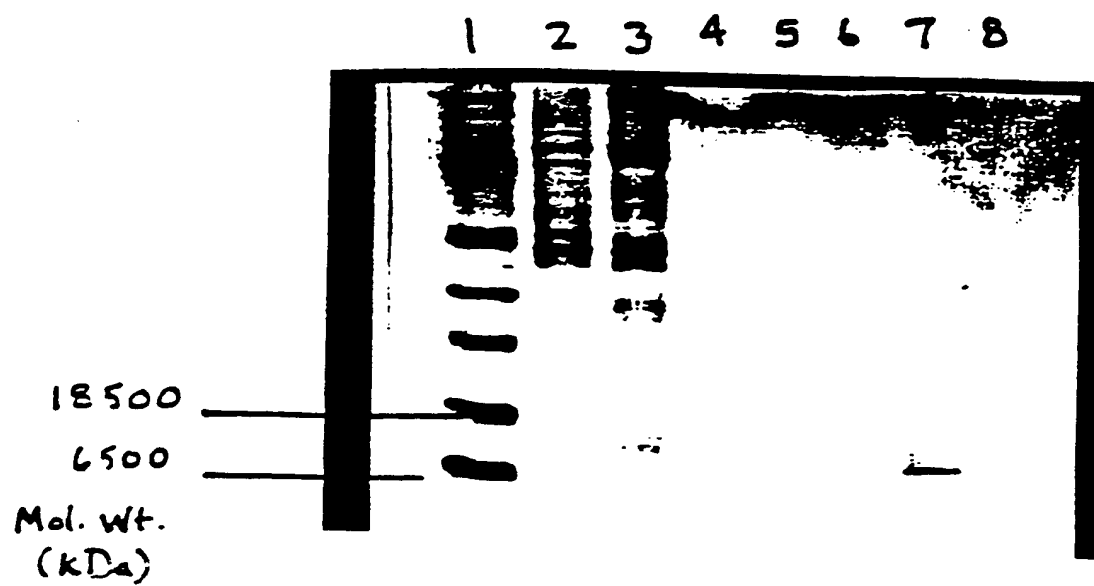


Fig 1



**Titration of wCRW Gut Proteases  
with E-64 and Virgiferin**

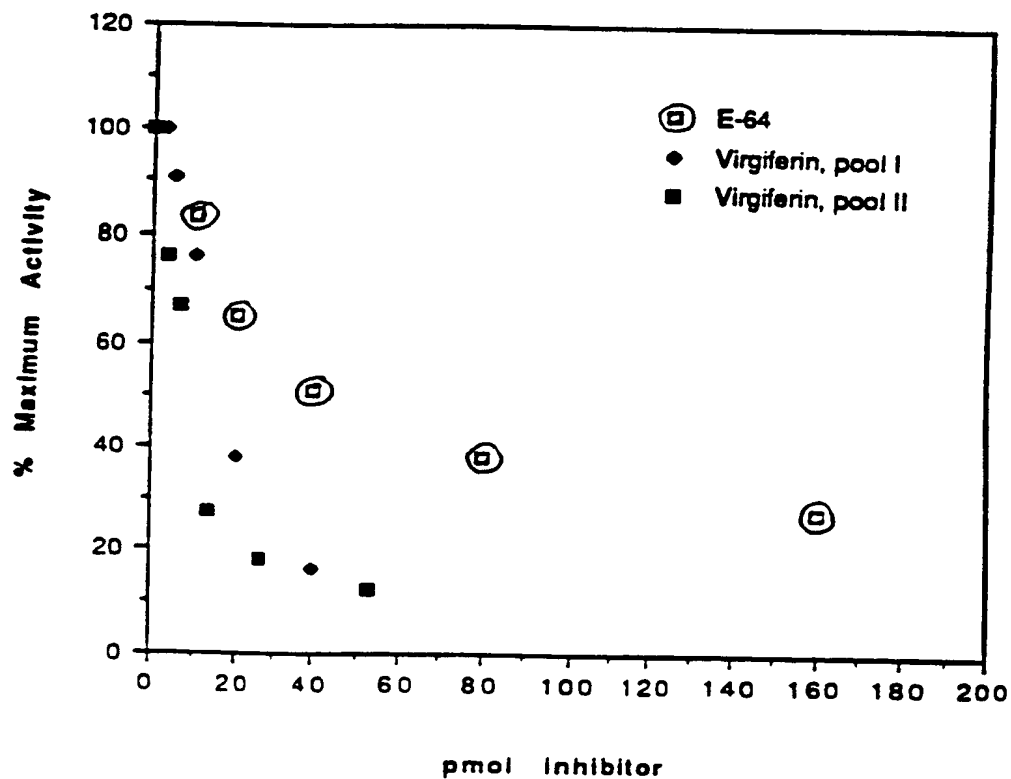


Fig 2

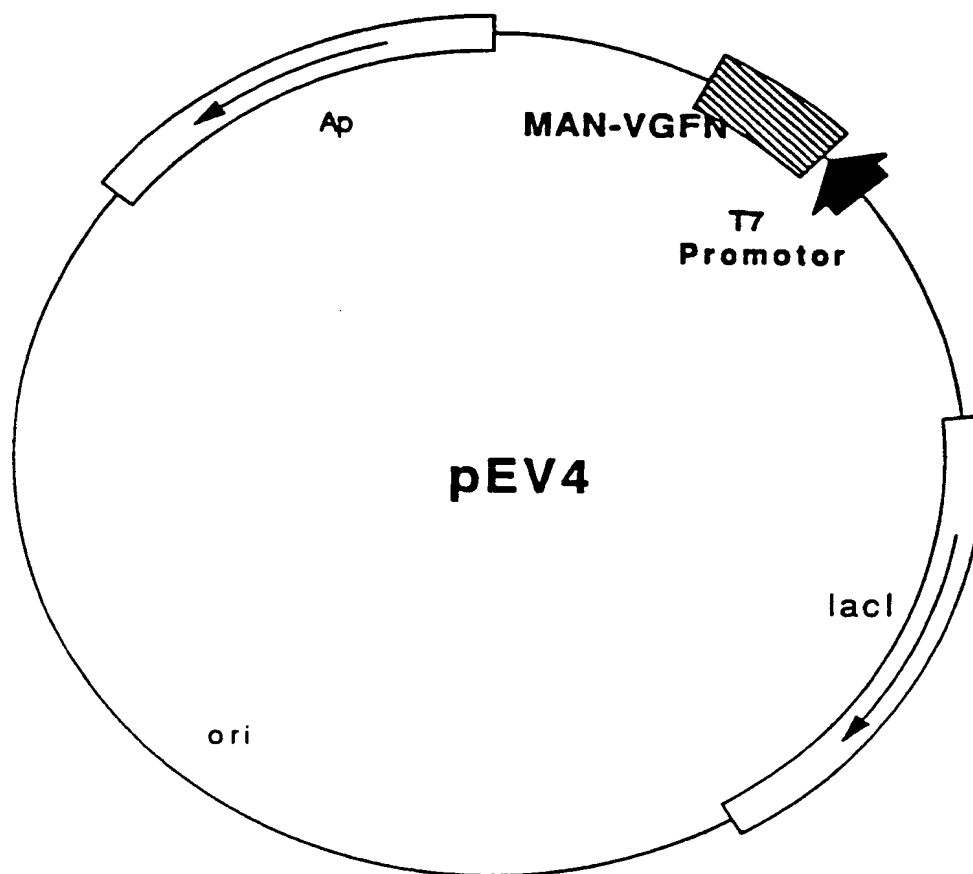


Fig 3

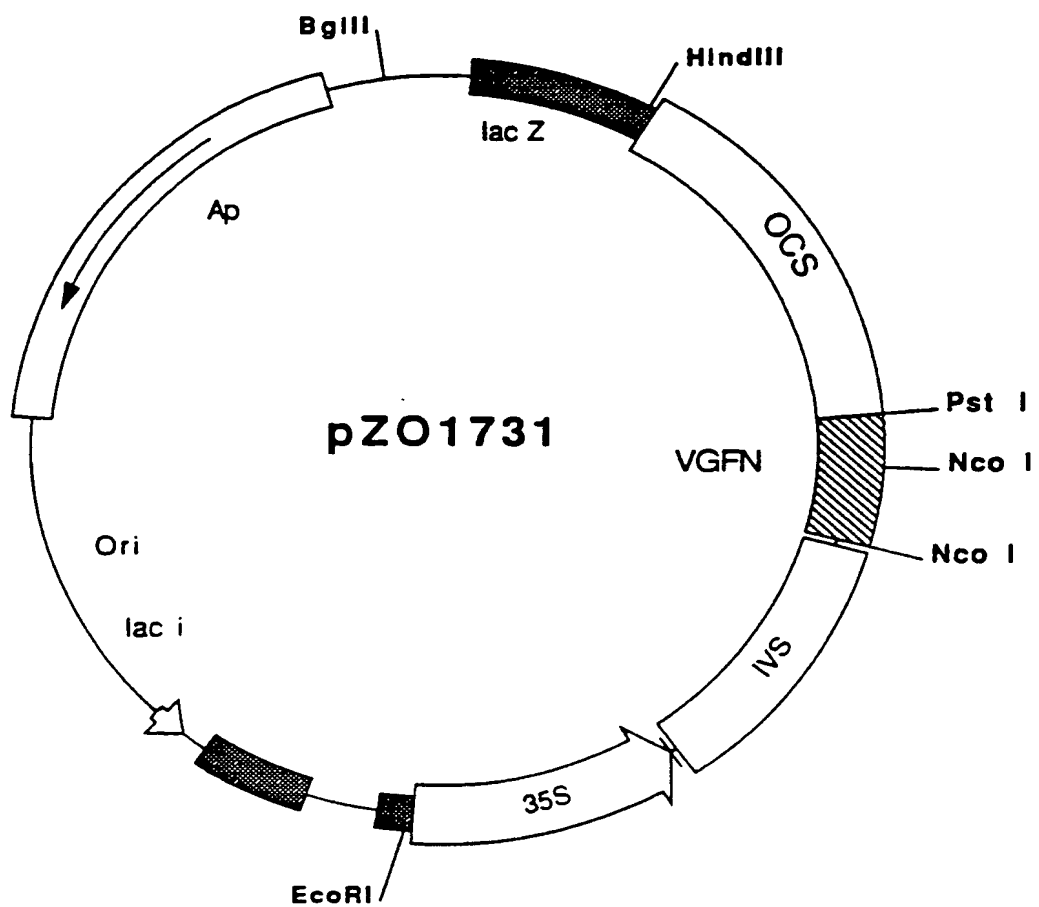


Fig 4

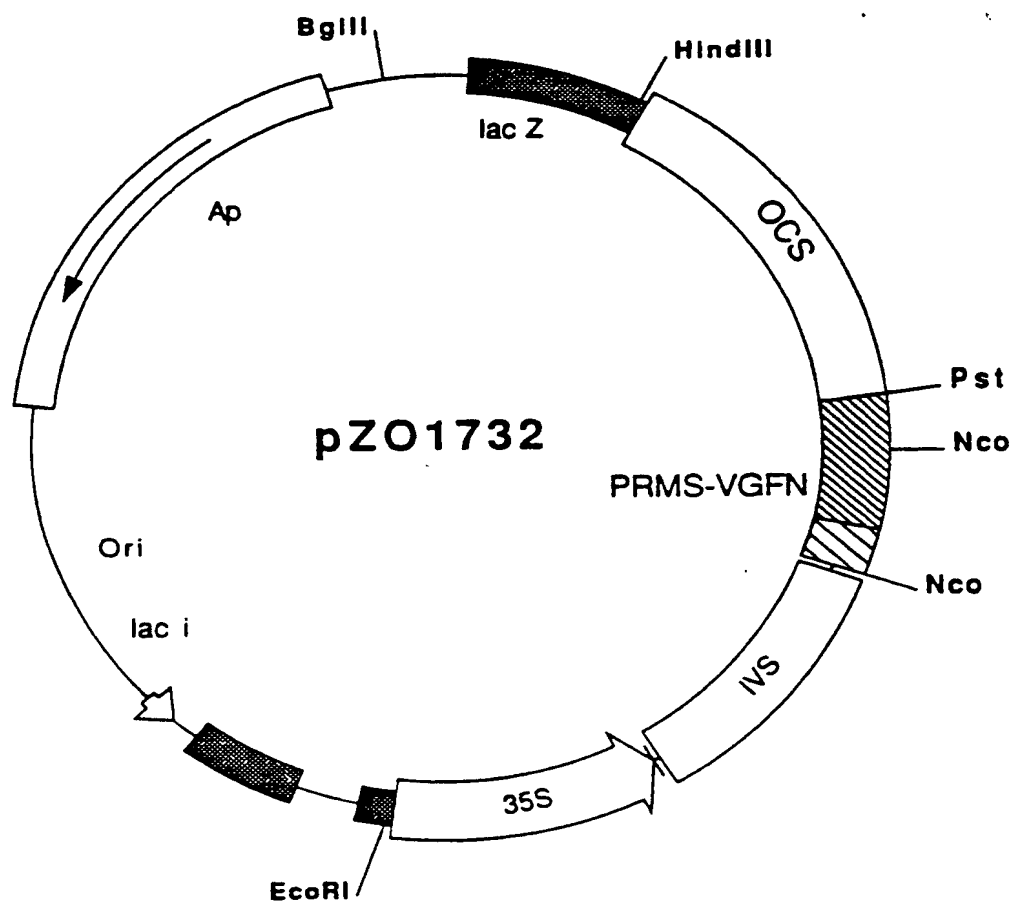


Fig 5

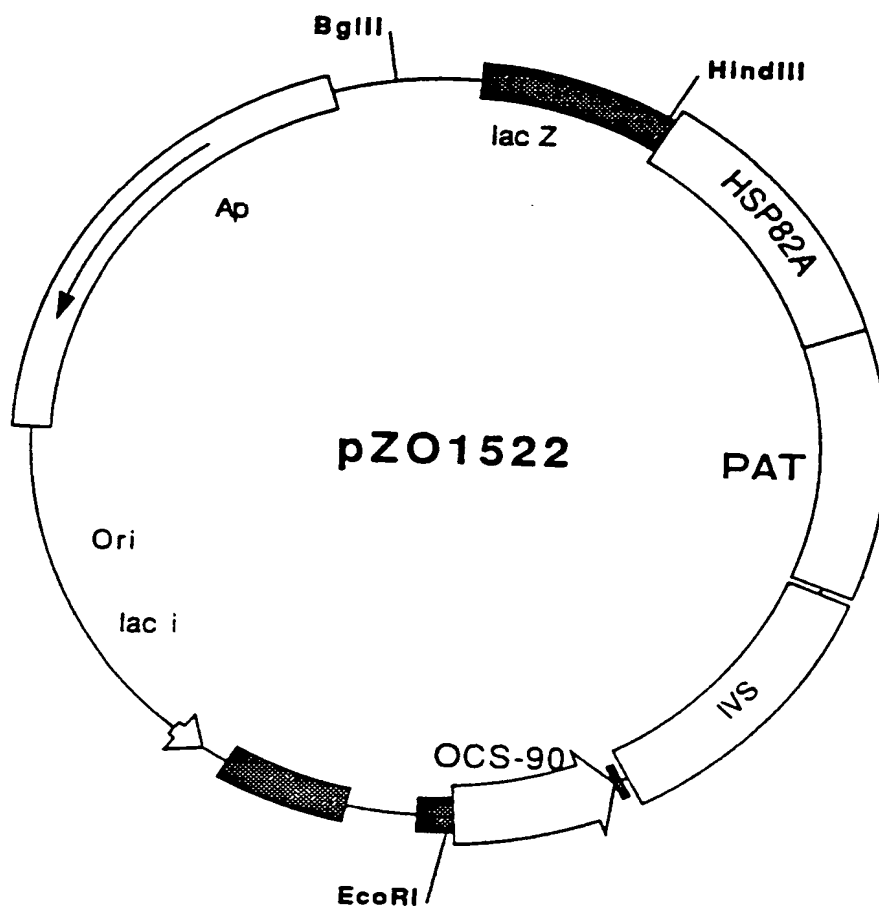


Fig 6

# INTERNATIONAL SEARCH REPORT

Int. Appl. No.  
PCT/EP 95/00881

A. CLASSIFICATION OF SUBJECT MATTER  
IPC 6 C12N15/12 C12N15/82 C07K14/81 A01N63/02 A01H5/00

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N C07K A01N A01H

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO-A-94 02504 (MILES INC ;MULLER DANIEL (US); DELARIA KATHERINE (US); WALLACE LIN) 3 February 1994 see the whole document ---	1-7
X	FEBS LETTERS, vol. 292, November 1991 AMSTERDAM NL, pages 115-120, LAYCOCK, M.V., ET AL. 'MOLECULAR CLONING OF THREE CDNAS THAT ENCODE CYSTEINE PROTEINASES IN THE DIGESTIVE GLAND OF AMERICAN LOBSTER (HOMARUS AMERICANUS)' see figure 2 --- -/--	4,5

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

### \* Special categories of cited documents :

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

- \*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- \*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- \*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- \*&\* document member of the same patent family

Date of the actual completion of the international search

14 June 1995

Date of mailing of the international search report

21-06-1995

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl,  
Fax (+ 31-70) 340-3016

Authorized officer

Maddox, A

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 95/00881

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	FEBS LETTERS, vol. 301, no. 1, April 1992 AMSTERDAM NL, page 125 LAYCOCK, M.V., ET AL. 'MOLECULAR CLONING OF THREE CDNAS THAT ENCODE CYSTEINE PROTEINASES IN THE DIGESTIVE GLAND OF AMERICAN LOBSTER (HOMARUS AMERICANUS)' see figure 2 ---	4,5
X	EUR. J. IMMUNOL., vol. 19, 1989 pages 631-635, DENIZOT, F., ET AL. 'NOVEL STRUCTURES CTLA-2ALPHA AND CTLA-2BETA EXPRESSED IN MOUSE ACTIVATED T CELLS AND MAST CELLS AND HOMOLOGOUS TO CSTEINE PROTEINASE PROREGIONS' see the whole document ---	4,5
X	PLANT MOLECULAR BIOLOGY, vol. 21, 1993 pages 655-663, MASOUD, S.A., ET AL. 'EXPRESSION OF A CYSTEINE PROTEINASE INHIBITOR (ORYZACYSTATIN-I) IN TRNASGENIC TOBACCO PLANTS' see the whole document ---	3-6,8
X	PROTEIN EXPRESSION AND PURIFICATION, vol. 3, 1992 pages 41-49, CHEN, M.-S., ET AL. 'RICE CYSTATIN: BACTERIAL EXPRESSION, PURIFICATION, CYSTEINE PROTEINASE INHIBITORY ACTIVITY, AND INSECT GROWTH SUPPRESSING ACTIVITY OF A TRUNCATED FORM OF THE PROTEIN' see the whole document ---	3-6,11
X	WO-A-92 21753 (DOWELANCO) 10 December 1992 see the whole document ---	3-6,8-11
X	EP-A-0 348 348 (CIBA GEIGY AG) 27 December 1989 see the whole document ---	3-6,8-11
X	DATABASE WPI Section Ch, Week 9342 Derwent Publications Ltd., London, GB; Class C03, AN 93-331338 & JP,A,05 238 913 (SUNTORY) , 17 September 1993 see abstract --- -/--	11

## C(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	J. CELL. BIOCHEM. SUPPL.18D. 1994, page 134 GUEGLER, K., ET AL. 'CLONING AND CHARACTERIZATION OF THIOL PROTEASES FROM THE FALSE COLORADO POTATO BEETLE' see abstract s115 & KEYSTONE SYMPOSIUM ON STRUCTURAL AND MOLECULAR BIOLOGY OF PROTEASE FUNCTION AND INHIBITION, HELD MARCH 5-12,1994., ---	1
A	BIOESSAYS, vol. 10, no. 1, January 1989 pages 20-24, RYAN, C.A. 'PROTEINASE INHIBITOR GENE FAMILIES: STRATEGIES FOR TRANSFORMATON TO IMPROVE PLANT DEFENSES AGAINST HERBIVOREA' see page 23, column 2 - column 3 ---	8-10
A	ANN. REV. PHYTOPATHOL., vol. 28, no. 9, 1990 pages 425-449, RYAN, C.A. 'PROTEINASE INHIBITORS IN PLANTS: GENES FOR IMPROVING DEFENSES AGAINST INSECTS AND PATHOGENS' see page 435 - page 437 ---	8-10
A	ABSTRACTS OF THE VIITH INTERNATIONAL CONGRESS ON PLANT TISSUE AND CELL CULTURE, HELD AMSTERDAM, JUNE 24-29, 1990., page 78 THOMAS, J.C., ET AL. 'INSECT PROTEINASE INHIBITOR EXPRESSION IN TRANSGENIC TOBACCO AND COTTON' see abstract A2-137 ---	8-10
A	ARCH INSECT BIOCHEM PHYSIOL 19 (4). 1992. 285-298., GILLIKIN J W, ET AL. 'PARTIAL CHARACTERIZATION OF DIGESTIVE TRACT PROTEINASES FROM WESTERN CORN ROOTWORM LARVAE DIABROTICA-VIRGIFERA' see the whole document -----	1-11



# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/EP 95/00881

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO-A-9402504	03-02-94	CA-A- 2140127 EP-A- 0651764	03-02-94 10-05-95
WO-A-9221753	10-12-92	AU-A- 2235692 BR-A- 9206118 EP-A- 0587798 JP-T- 6508033	08-01-93 27-12-94 23-03-94 14-09-94
EP-A-0348348	27-12-89	AU-B- 631551 AU-A- 3656889 JP-A- 2046238	03-12-92 21-12-89 15-02-90

